ВЕЙВЛЕТ-АНАЛИЗ В ИССЛЕДОВАНИИ НЕОДНОРОДНОСТЕЙ СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

Лобач В. И.
Кафедра математического моделирования и анализа данных,
Белорусский государственный университет
Минск, Республика Беларусь
E-mail: lobach@bsu.by

Предлагается и исследуется статистический критерий обнаружения неоднородностей случайных последовательностей на основе вейвлета Хаара.

Введение

Информация и изучение ее свойств представляют огромную важность в мире современных компьютерных технологий. Все более широкое распространение при обработке информационных массивов данных приобретают методы вейвлет-анализа [1–3]. Вейвлеты в последнее десятилетие нашли также широкие применения в обработке сигналов и изображений. Теория вейвлетов является альтернативной анализу Фурье и дает более гибкую технику обработки сигналов. Одно из основных преимуществ вейвлет-анализа заключается в том, что он позволяет заметить хорошо локализованные свойства сигнала. Анализ Фурье этого не дает, так как в коэффициентах Фурье отражается поведение сигнала за все время его существоввания. Целью данной работы является обнаружение неоднородностей случайных последовательностей при помощи данных методов.

I. Вейвлет Хаара

Определение 1. *Функция $\psi \in L^2(R)$ называется ортогональным вейвлетом, если семейство $\{\psi_{j,k}\}$, определенное формулой $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$, $j,k \in Z$, является ортонормированным базисом в $L^2(R)$; это означает, что*

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \cdot \delta_{k,m}, \quad j,k,l,m \in Z,$$

и любая функция $f \in L^2(R)$ может быть представлена как

$$f(t) = \sum_{j,k=-\infty}^{\infty} c_{j,k}\psi_{j,k}(t), \tag{1}$$

где ряд (1) сходится в $L^2(R)$, а именно

$$\lim_{M_1,N_1,M_2,N_2\to\infty} \left\| f - \sum_{j=-M_2}^{N_2} \sum_{j=-M_1}^{N_1} c_{j,k}\psi_{j,k} \right\|_2 = 0.$$

Ряды, представляющие функции f в (1), называются *вейвлет-рядами.* Вейвлет-коэффициенты $c_{j,k}$ определяются формулой

$$c_{j,k} = \langle f \psi_{j,k} \rangle.$$

Определим интегральное преобразование W_ψ в $L^2(R)$ как

$$(W_\psi f)(b,a) = |a|^{-1/2} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt, \ f \in L^2(R),$$

то вейвлет-коэффициенты в (2) принимают вид

$$c_{j,k} = (W_\psi f)\left(\frac{k}{2^j}, \frac{1}{2^j}\right).$$

Рассматривается случайная последовательность:

$$\Xi = \left\{\xi_t | \xi_t \in \{0,1\}, \, t = \overline{0, N-1}\right\} \tag{2}$$

Положим, не теряя общности, $N = 2^n$, где $n \in N$.

Нулевая гипотеза о распределении вероятностей данной последовательности формулируется в виде:
$H_0 : \{\xi_t\}$ – последовательность случайных, независимых и одинаково распределенных величин.
Альтернативная гипотеза:
$H_1 : \xi_t = \eta_t \oplus a_t$, $t = \overline{0, 2^n - 1}$, где $A = \left\{a_t | a_t \in \{0,1\}, \, t = \overline{0, 2^n - 1}\right\}$ – переодическая с периодом T последовательность детерминированных величин;
$G = \left\{\eta_t | \eta_t \in \{0,1\}, \, t = \overline{0, 2^n - 1}\right\}$ – последовательность независимых случайных величин, распределенных по закону Бернулли $Bi(1,p)$, причем $p \ll 1/2$.

Определение 2. *Дискретное вейвлет-преобразование случайной последовательности (2) задается путем вычисления коэффициентов:*

$$d_{j,k}^{(\psi)} = \sum_{t=0}^{2^n-1} \xi_t \psi_{j,k}(t).$$

В качестве вейвлет-функции выберем вейвлет Хаара :

$$\psi^H(t) = \begin{cases} 1, & 0 \le t < \frac{1}{2}, \\ -1, & \frac{1}{2} \le t < 1, \\ 0, & \text{иначе.} \end{cases}$$

Для последовательности (2) и дискретного вейвлет-преобразования Хаара детализирующие (уточняющие) коэффициенты преобразования имеют дискретное распределение вероятностей. Найдем это распределение вероятностей.

Ортонормированная система вейвлетов Хаара выглядит следующим образом:

$$\psi_{j,k}^{H}(t)=\begin{cases} 2^{-\frac{j}{2}}, & 2^j k\le t< \left(2^j k+2^{j-1}\right), \\ -2^{-\frac{j}{2}}, & \left(2^j k+2^{j-1}\right)\le t<\left(2^j k+2^j\right), \\ 0, & \text{иначе.} \end{cases}$$

Можно показать, что множеством значений $d_{j,k}^{(\psi)}$ является $\pm 2^{\frac{j}{2}-1}$; $\pm\left(2^{\frac{j}{2}-1}-2^{-\frac{j}{2}}\right)$; $\pm\left(2^{\frac{j}{2}-1}-2\cdot 2^{-\frac{j}{2}}\right)$; $\ldots$; $\pm\left(2^{\frac{j}{2}-1}-(2^{j-1}-1)\cdot 2^{-\frac{j}{2}}\right)$; 0.

Т. е. окончательно можно записать:

$$d_{j,k}^{(\psi)}=\pm\left(2^{\frac{j}{2}-1}-i\cdot 2^{-\frac{j}{2}}\right),\ i=\overline{0,2^{j-1}},\ j=\overline{1,n}.$$

II. Результаты компьютерных экспериментов

Приведем теорему, которая отражает свойства вейвлет-коэффициентов и является основой для построения статистического χ^2-критерия проверки гипотезы об однородности исследуемой случайной последовательности.

Теорема [4]. *Пусть элементы последовательности (2)* $\Xi=\left\{\xi_t|\xi_t\in\{0,1\},\ t=\overline{0,2^n-1}\right\}$ *независимы и одинаково распределены* $\mathrm{P}\{\xi_t{=}0\}{=}\mathrm{P}\{\xi_t{=}1\}{=}j/2$, *тогда детализирующие коэффициенты имеют следующее распределение*

$$\mathrm{P}\left\{d_{j,k}^{(\psi)}=\pm\left(2^{\frac{j}{2}-1}-j\cdot 2^{-\frac{j}{2}}\right)\right\}=$$

$$=\frac{\sum\limits_{K=0}^{i}C_{2^{j-1}}^{K}C_{2^{j-1}}^{i-K}}{2^{2^t}},$$

$$i=\overline{0,2^{j-1}},\ j=\overline{1,n}.$$

На основе этой теоремы построен критерий проверки гипотезы H_0, он использует χ^2-статистику Пирсона для проверки гипотезы соответствия эмпирического и теоретического распределений вероятностей детализирующих коэффициентов. Проверка проводилась для каждого уровня детализации $j=\overline{1,n}$, гипотеза H_0 принимается, если на всех уровнях принималась нулевая гипотеза χ^2-критерия.

Методом статистического моделирования были получены оценки вероятностей ошибок первого и второго рода. Оценка вероятности ошибки первого рода составила 0,045 при уровне значимости $\alpha{=}0,05$. Для оценки вероятности ошибки второго рода β моделировалась случайная последовательность с вероятностью $p{=}\mathrm{P}\{\xi_t{=}1\}$, которая изменялась от 0 до 0,45. Результаты моделирования приведены в таблице.

p	0	0,05	0,01	0,1	0,2	0,3	0,4	0,45
β	0	0,01	0,02	0,25	0,028	0,032	0,2	0,54

Результаты моделирования показывают эффективность применения вейвлета Хаара в задаче обнаружения неоднородностей в дискретных случайных последовательностях, которые являются основными математическими моделями дискретных сигналов.

1. Chiann, Ch. A wavelet analysis for time series / Ch. Chiann, P. A. Morretin // Journal of Nonparametric Statistics. – 1998. –Vol. 10, № 1. – P. 1–46.
2. Астафьева, Н. М. Вейвлет-анализ: основы теории и примеры применения /Н. М. Астафьева // Успехи физических наук. – 1996. No. 11. – С. 1145–1170.
3. Чуи, К. Введение в вейвлеты / К. Чуи // М.: Мир. – 2001, 285 с.
4. Андрухаев, Х. М. Сборник задач по вероятностей / Х. М. Андрухаев // М.: Просвещение. – 1985, 160 с.